CONTRUCCIÓN DE UN CLUSTER DE COMPUTADORAS

Ing. Ricardo Coloma

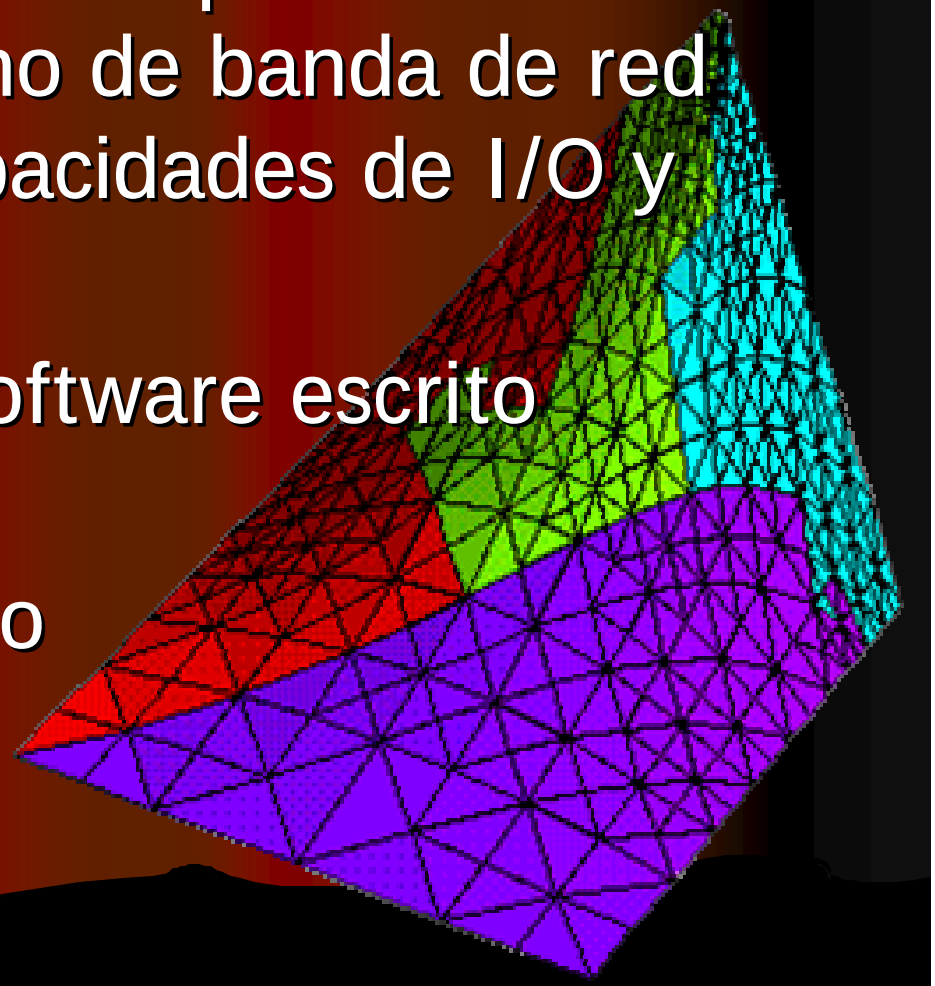
ROJ

INDICE

- Introducción
- Arquitecturas y porqué Cluster
- Componentes del Hardware y Software
- Soluciones con Redes
- Administración del Cluster
- El Nodo en instalación y nuevos Nodos
- Programación y Librerías
- Ambiente Paralelo

Introducción

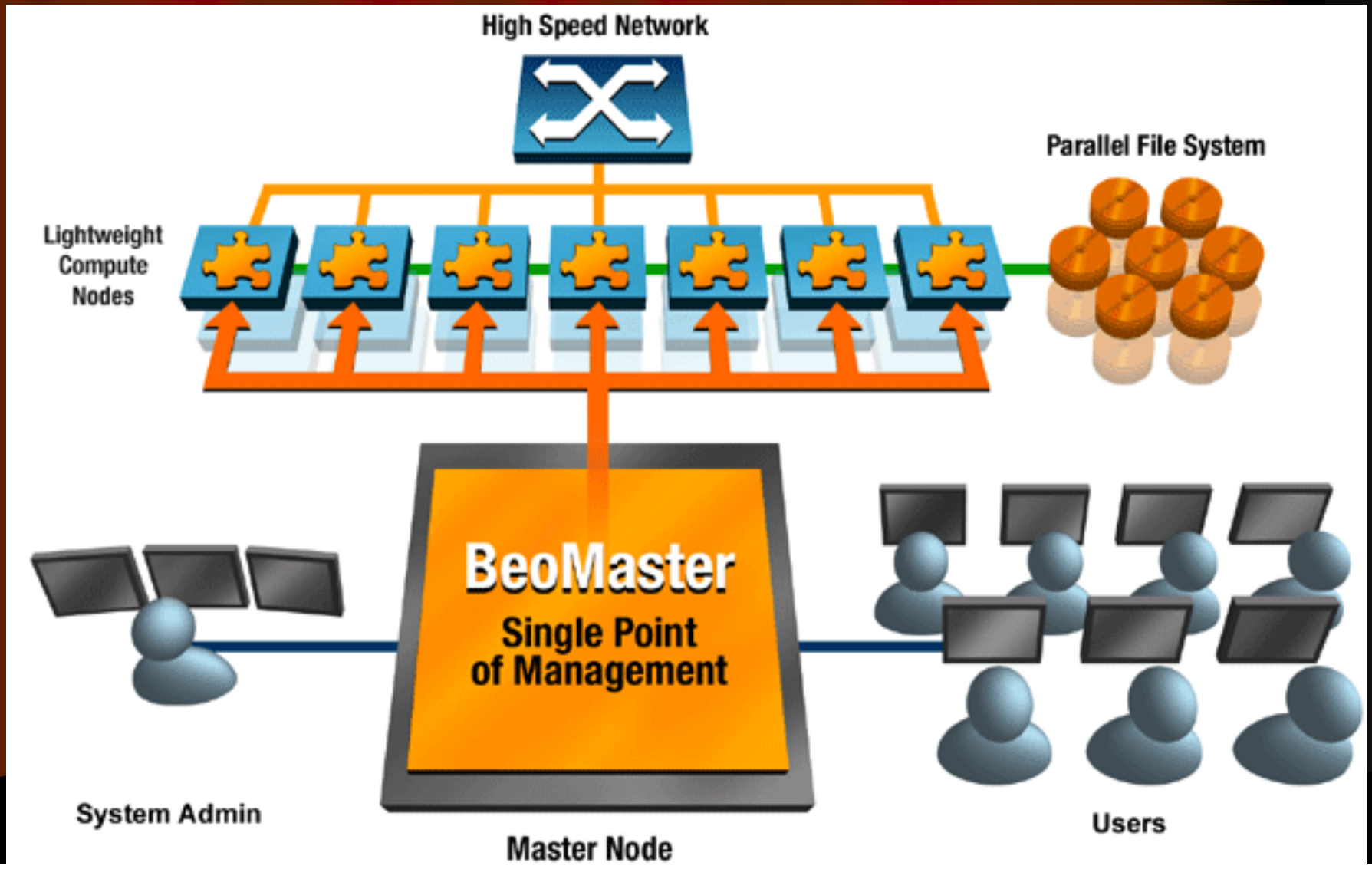
- Supercomputadoras= Tiempo de procesamiento + ancho de banda de red interna + grandes capacidades de I/O y almacenamiento
- Sistema operativo y software escrito especialmente
- Procesamiento Paralelo
- No ampliable



Arquitecturas y porqué Cluster

- Procesamiento Paralelo
- El uso de “Array Processors” VAX Cluster
- Computo distribuido en base a MPI (message passing interfaces)
- Files Servers
- Cluster de PC en
- base a Linux
- Fácil Integración

Estructura


High Speed Network
Parallel File System
Lightweight
Compute
Nodes
BeoMaster
Single Point
of Management
System Admin
Users
Master Node

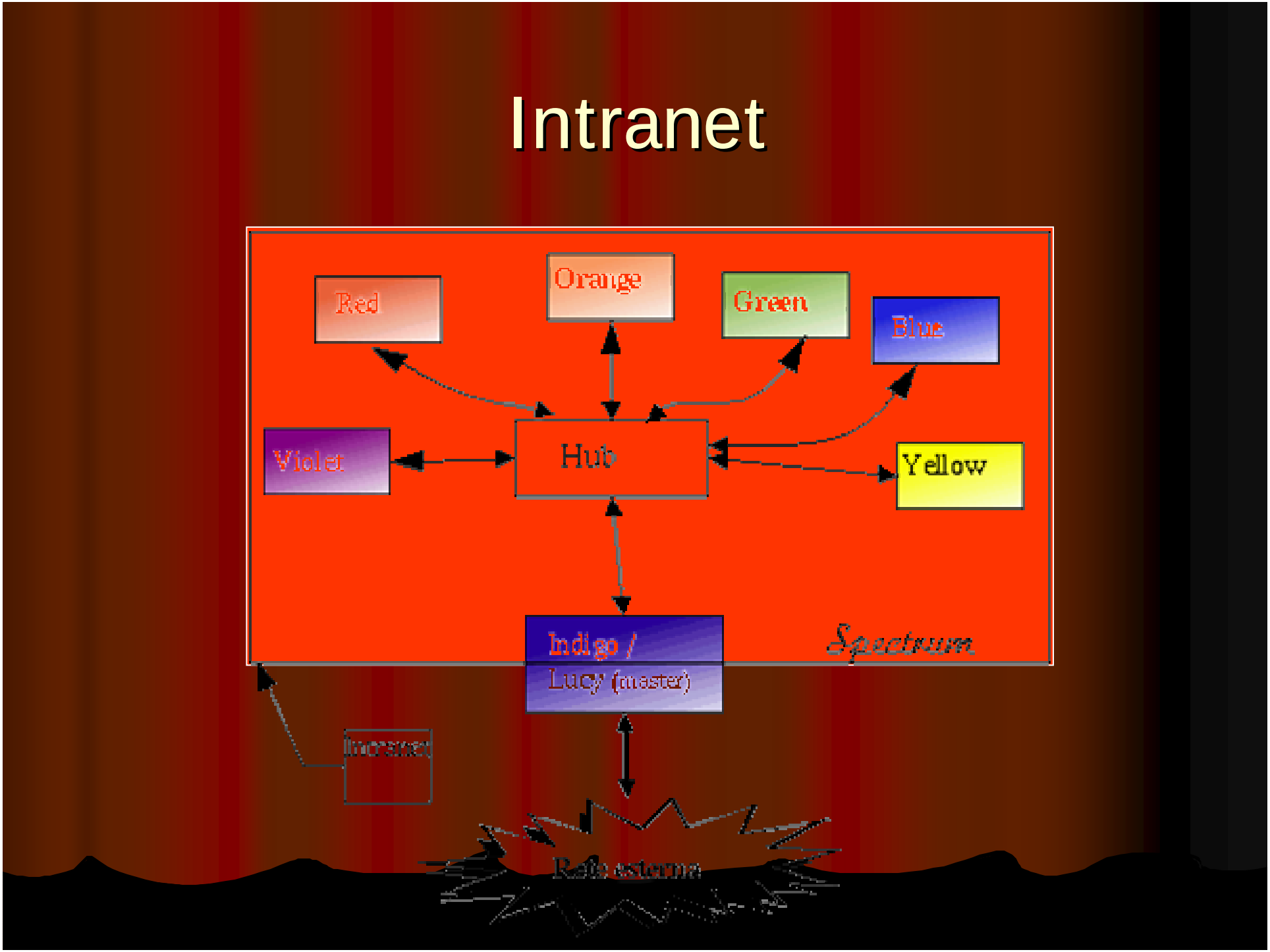
Hardware y Cluster

- CPUs, tecnologías RISC
- Paralelismo
  - Fino
  - Grueso
- Memorias
  - Cache - nanosegundos
  - Principal – 100 nsec
  - DDR
- Velocidad del Buss
- I/O
- Disco Duro
  - RAID(Redundant Array of Inexpensive Disks)
  - SCSI

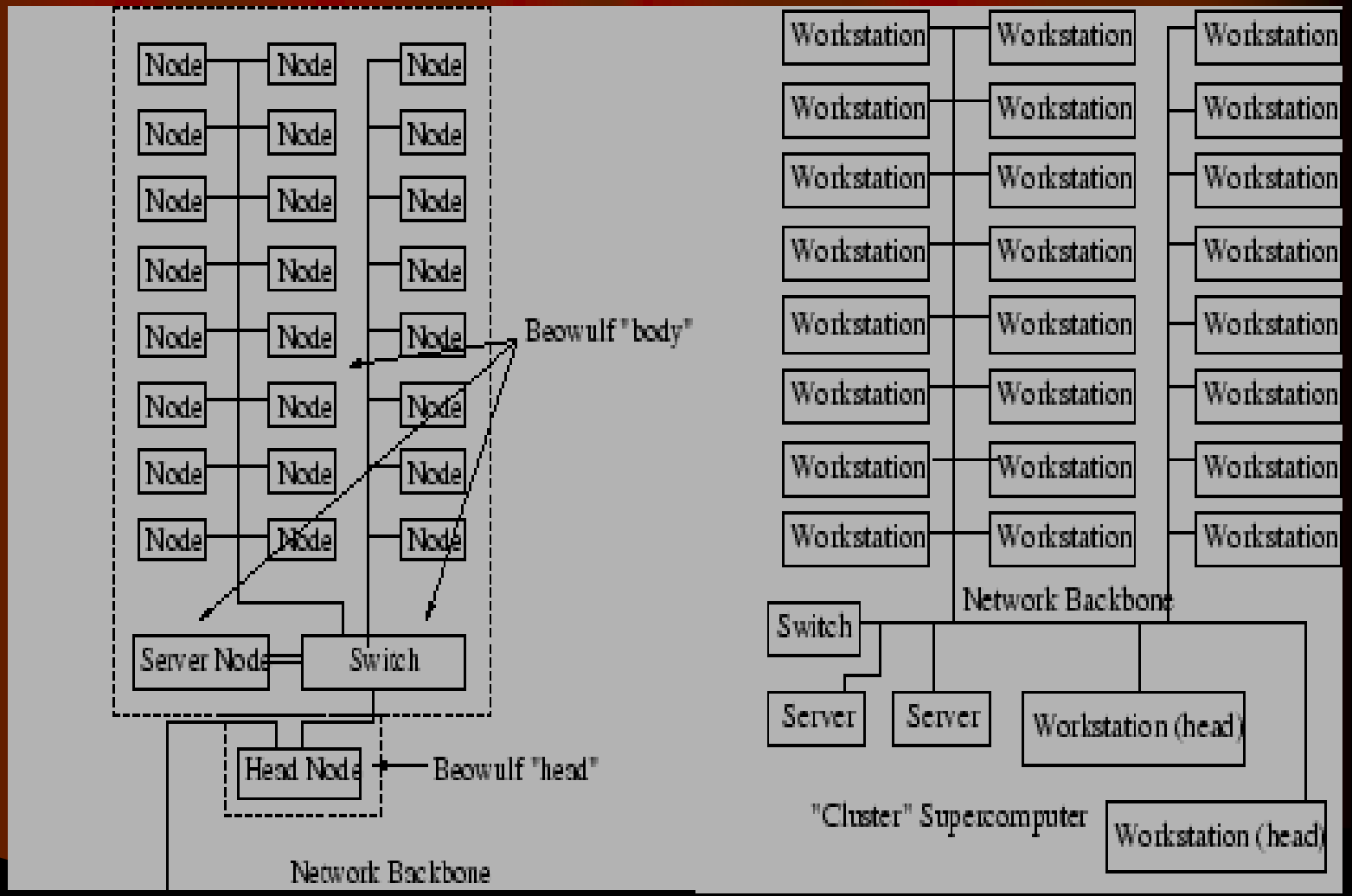
Software y Cluster

- Granularidad
- Dependencia
  - De los datos
  - Del control
- Multiprocesamiento
- Sincronización
- Subordinación de procesos hijos

Intranet


Red
Orange
Green
Blue
Violet
Hub
Yellow
Indigo /
Lucy (master)
Spectrum
Internet
Rete esterna

Cluster


Node
Server Node
Switch
Beowulf "body"
Head Node
Beowulf "head"
Network Backbone
Workstation
Switch
Server
Server
Workstation (head)
Network Backbone
"Cluster" Supercomputer
Workstation (head)

Configuración Simple

CLUSTER HOMOGENEO

(4) MICROCOMPUTADORAS

(1) KVM Switch= Keyboard, Video, Mouse

(1) LINUX RED HAT

CONFIGURACION INICIAL

MONITORES OPCIONALES

SOLO SE REQUIERE UNO , SI

SE DISPONE DEL MÓDULO

KVM X 4

CONFIGURACIÓN X 8

CLUSTER AMPLIADO A 8

MICROCOMPUTADORAS

PUEDE UTILIZARSE PLACAS

MADRE CON DOS PROCESADORES

XEON.

CLUSTER DE 16 PC

CLUSTER DE 64 PC

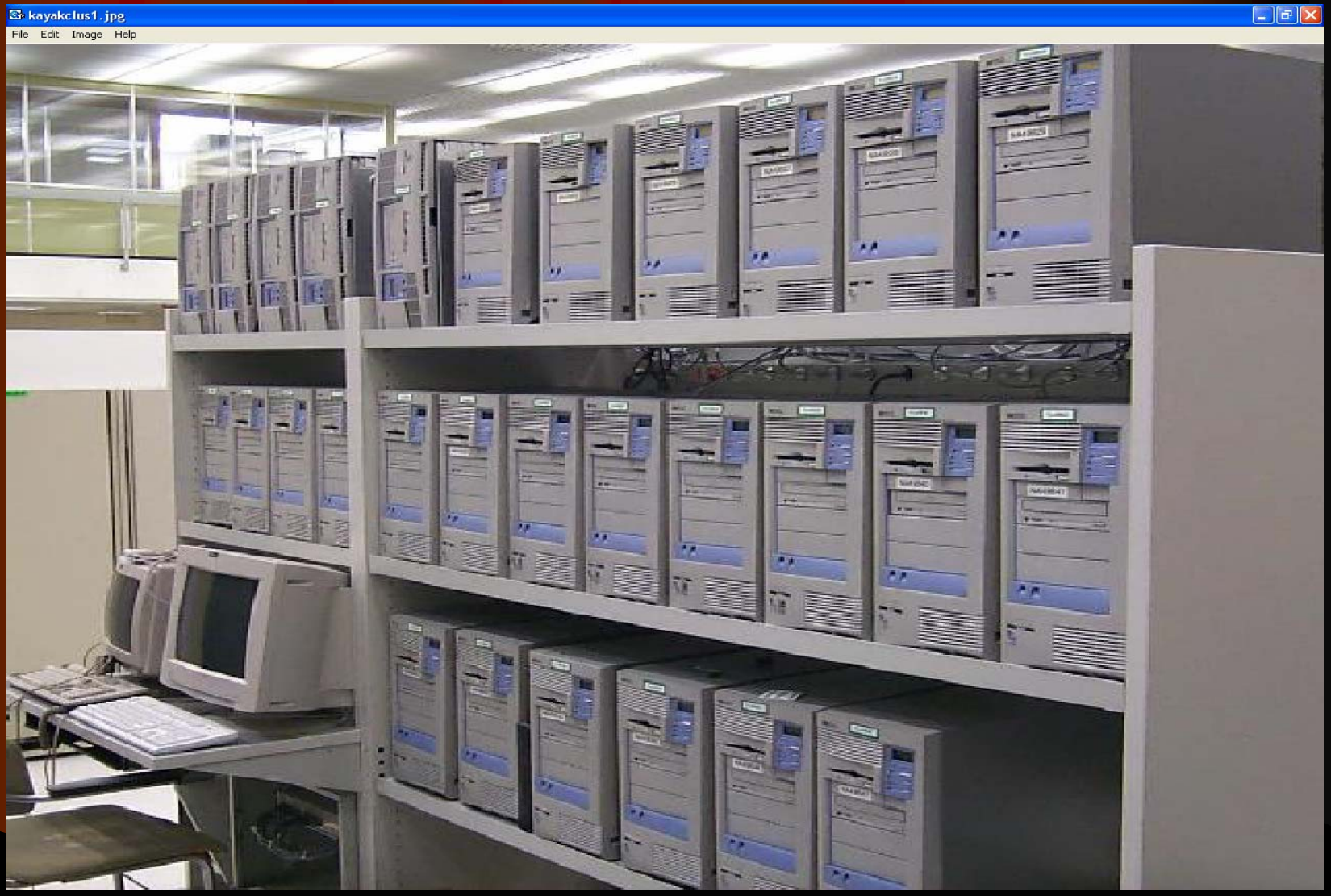
COMBINACION DE DIFERENTES

MODELOS, VERSIONES, CAPACIDADES

EN PROCESADORES, RAM, DISCO DURO

CASE Norma 3U (5.25 inch)
Norma 2U (3.5 inch)

CLUSTER 29 PC



CLUSTER MULTI-EQUIPO

CLUSTER HETEROGENEO

INTEGRACIÓN DE VARIADOS EQUIPOS MARCAS, MODELOS VELOCIDADES, RAM PROCESADORES, DISCOS ETC.

UPS

SUPER CLUSTER 64 PC

SISTEMAS OPERATIVOS
Kernel Linux

SIMPLE O MULTIPROCESADORES POR NODO

- REDHAT Linux
- Caldera OpenLinux
- SuSE Linux
- GNU/Debian Linux
- Slackware Linux
- TurboLinux
- SGI Linux

Componentes del Hardware y Software

- Solución que se busca encontrar
- Paralelismo y multi-procesamiento
  1. SMP- (Symmetric Multiprocessor)
  2. NUMA-(Non-Uniform Memory Access)
  3. UMA-(Uniform Memory Access)
  4. SIMD-(Single Instruction Multiple Data)
  5. MIMD-(Multiple Instruction Multiple Data, Linux Cluster por definición

Soluciones con Redes

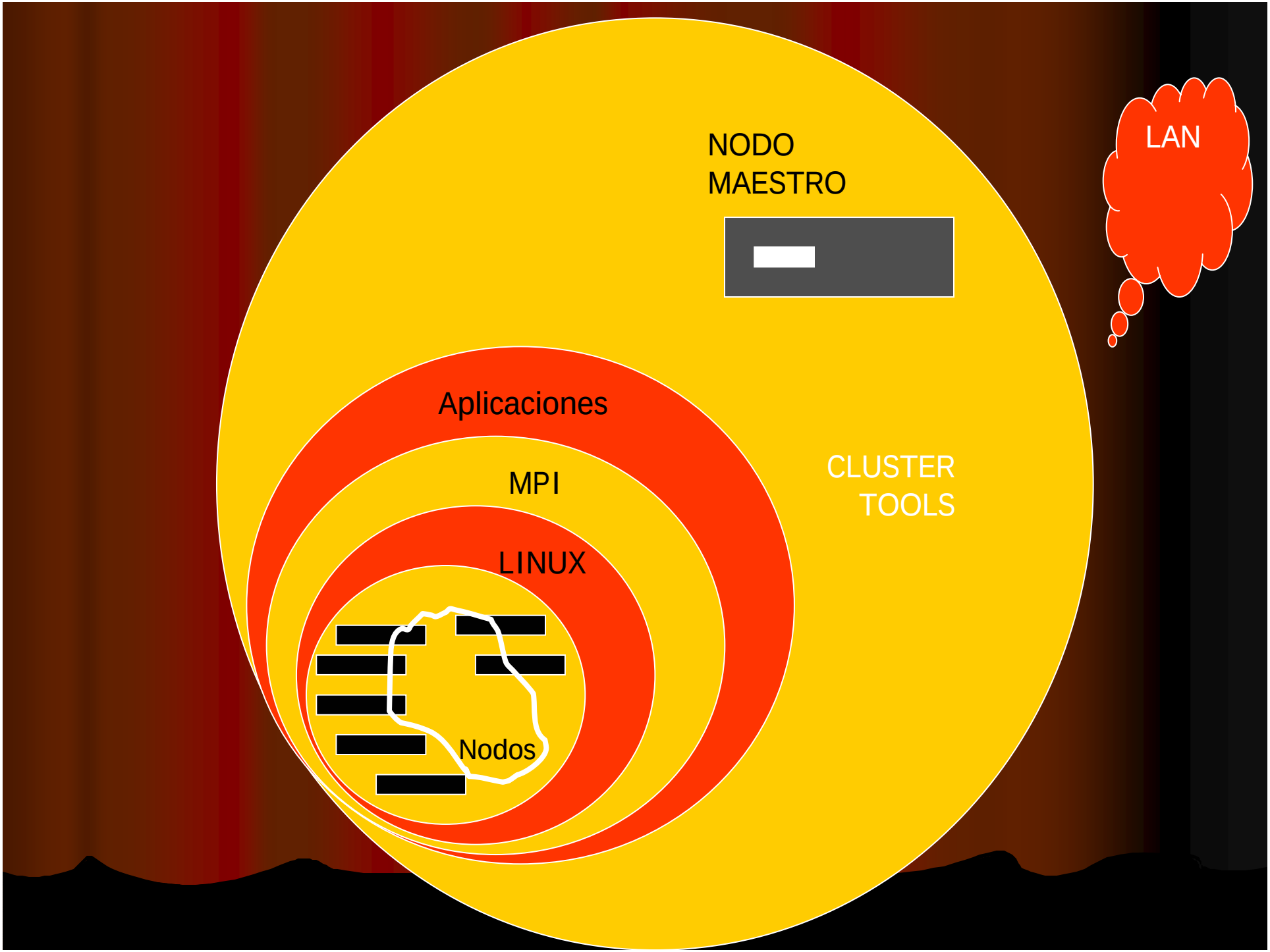
- Protocolos de redes Ethernet
- Interfaces del hardware (cable o fibra)
- Ancho de Banda en Mbits/seg.
- Hub , Routers o Switches
- Cubos , Hipercubos
- o Redes
- TCP/IP

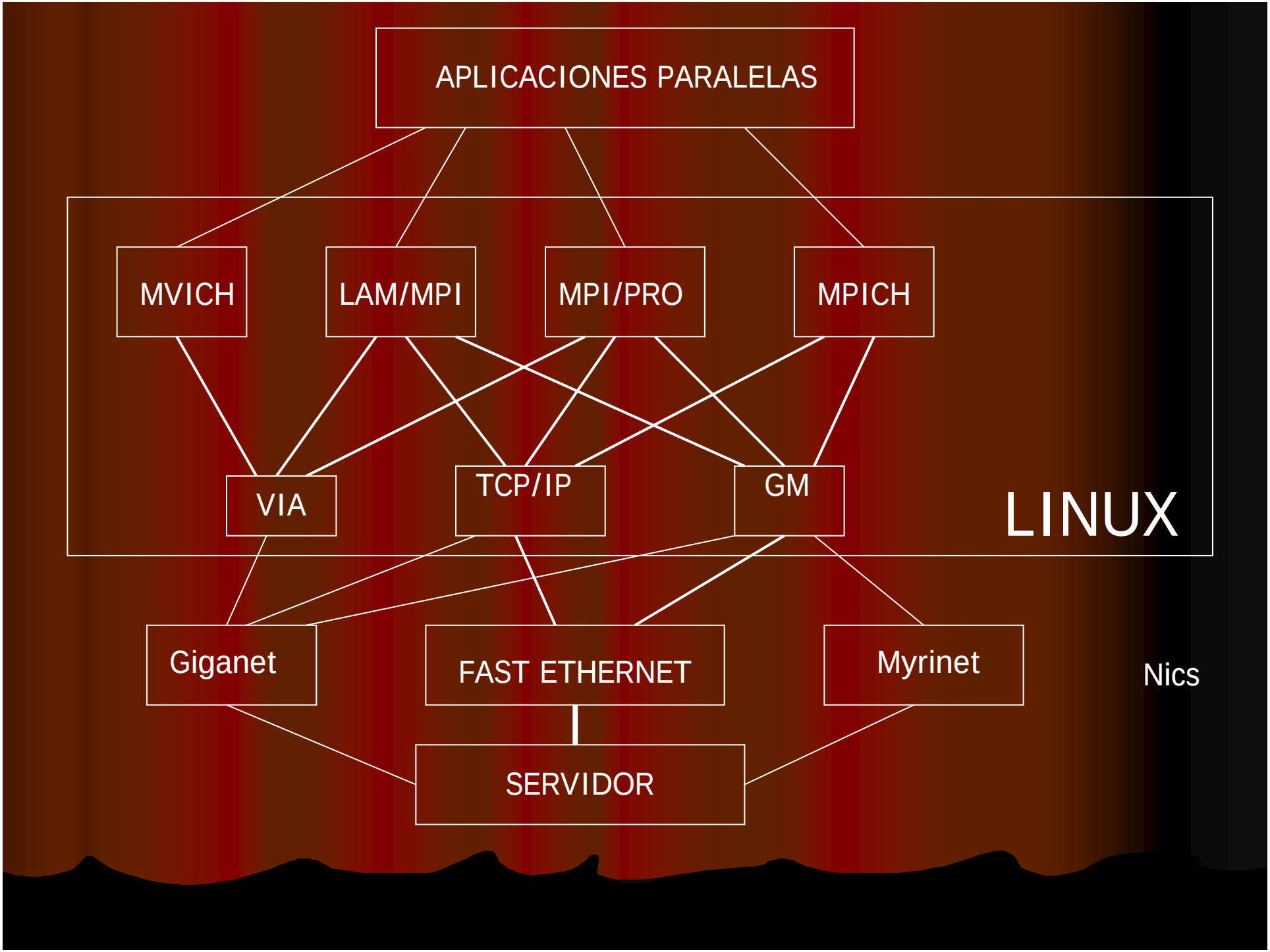
Comunicación Ethernet

- TCP/IP
  - Clase A 10.0.0.0
  - Clase B 172.16.0.0 - 172.31.0.0
  - Clase C 192.168.0.0 – 192.168.0.0
- Cadenas IP
- Seguridad

Sistemas de Programación

- MPI - Message Passing Interface
- PVM - Parallel Virtual Machine


NODO
MAESTRO
LAN
Aplicaciones
MPI
LINUX
Nodos
CLUSTER
TOOLS

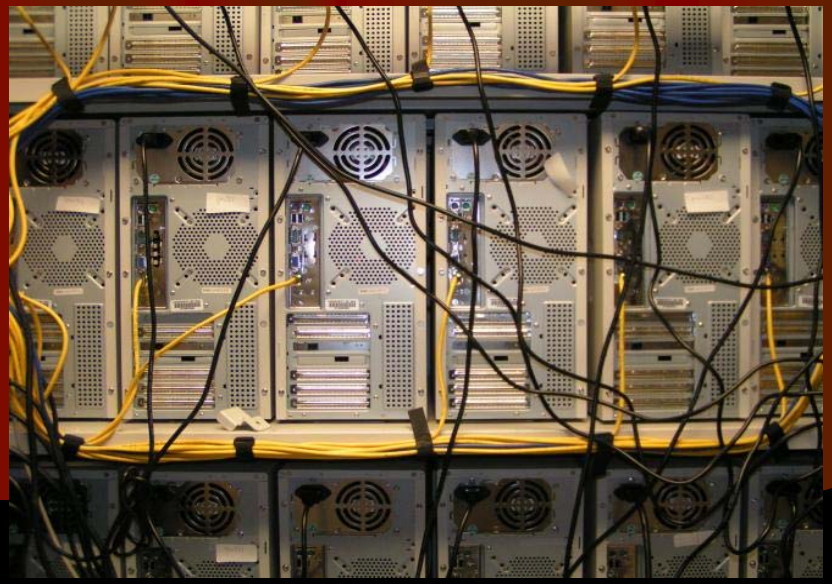

APLICACIONES PARALELAS
MVICH
LAM/MPI
MPI/PRO
MPICH
VIA
TCP/IP
GM
LINUX
Giganet
FAST ETHERNET
Myrinet
Nics
SERVIDOR

Configuración del Cluster

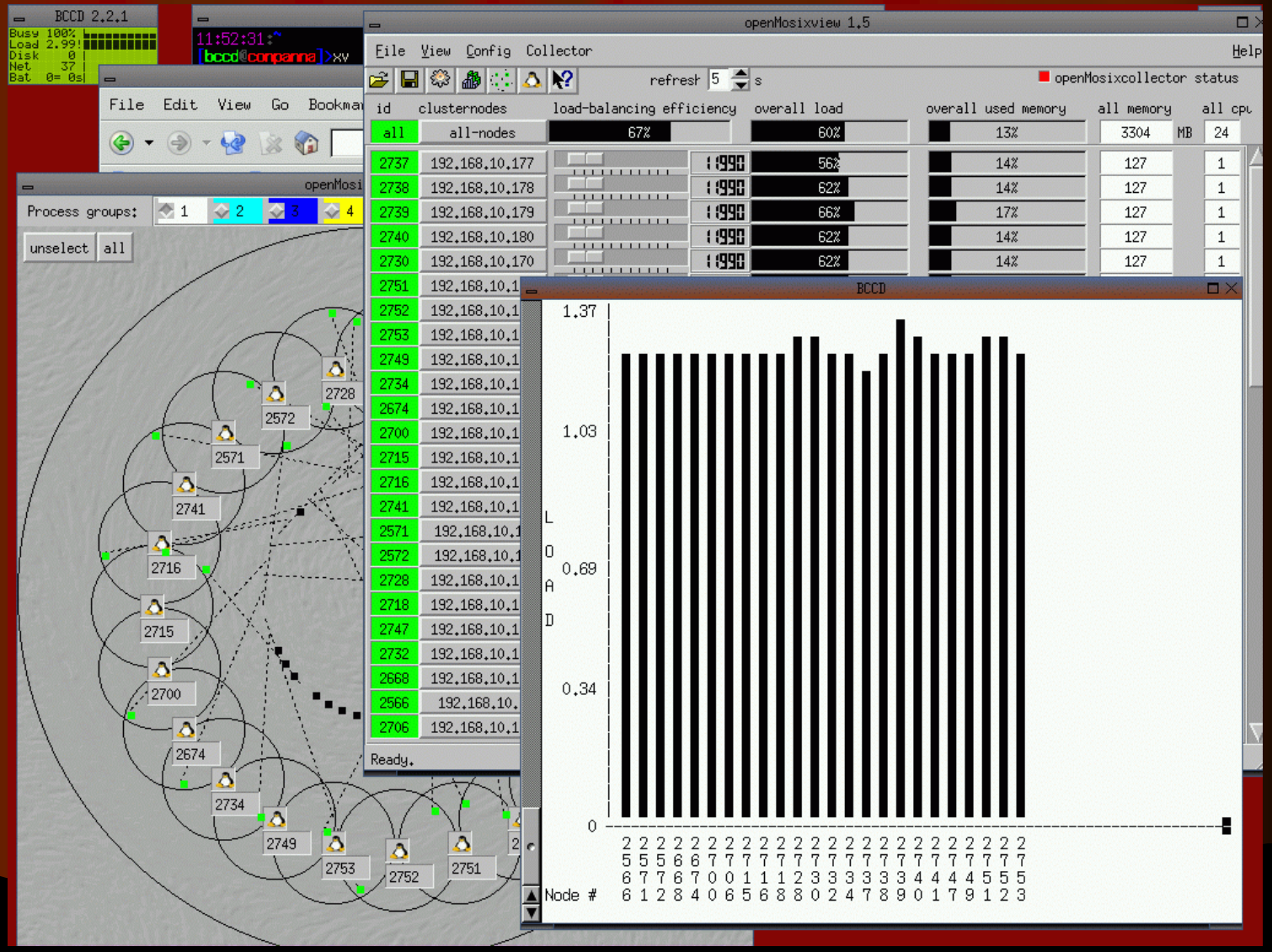
- Cubos o Hipercubos
- Mallas o Anillos
- Escalabilidad
- Acceso al Cluster
- Sistema de Mensajería

El Nodo en instalación y nuevos Nodos

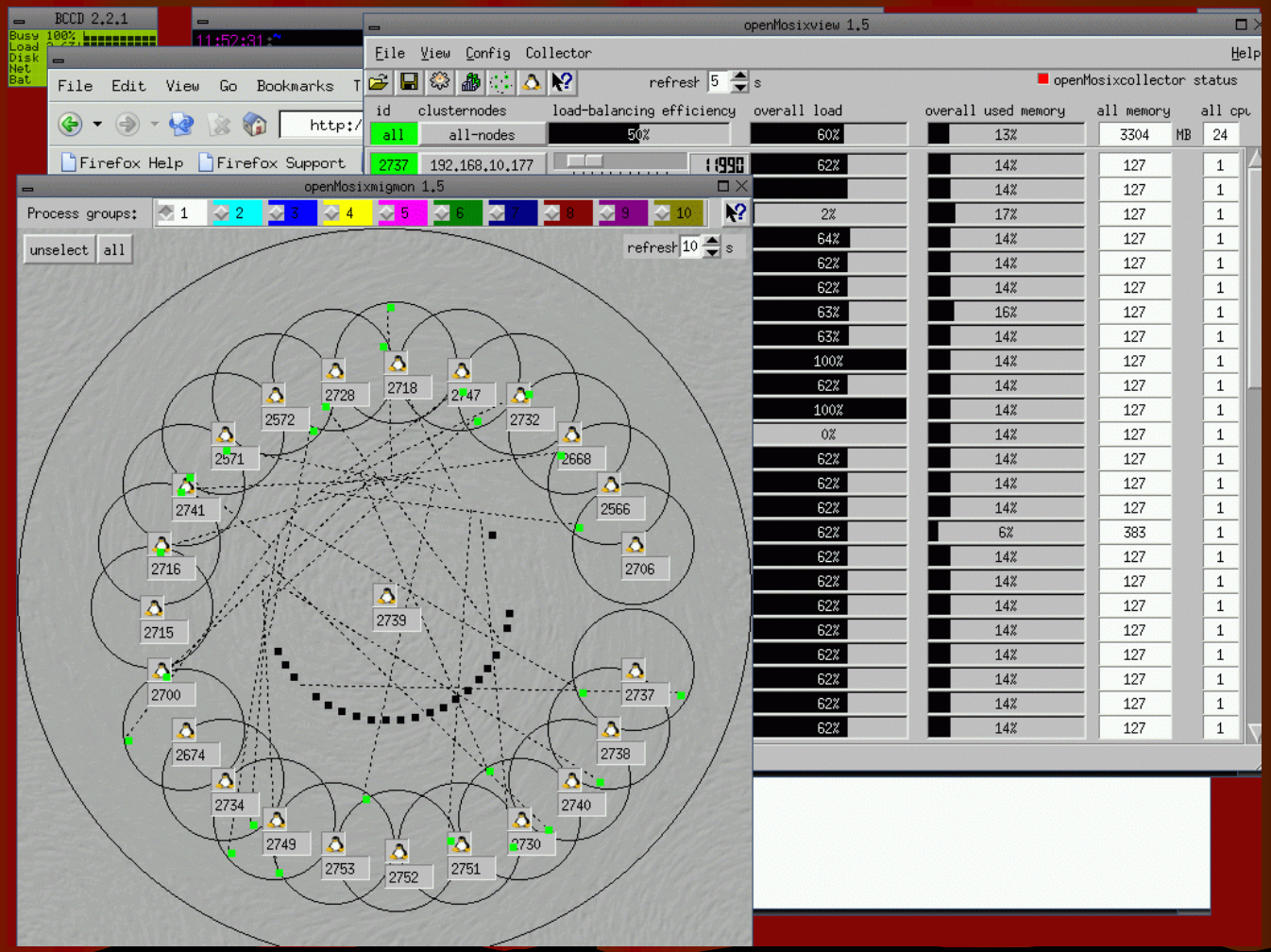
- El Nodo Maestro y el montaje del software
- El Boot del server
- Construcción de los nodos esclavos
- Administración de las cuentas y del sistema



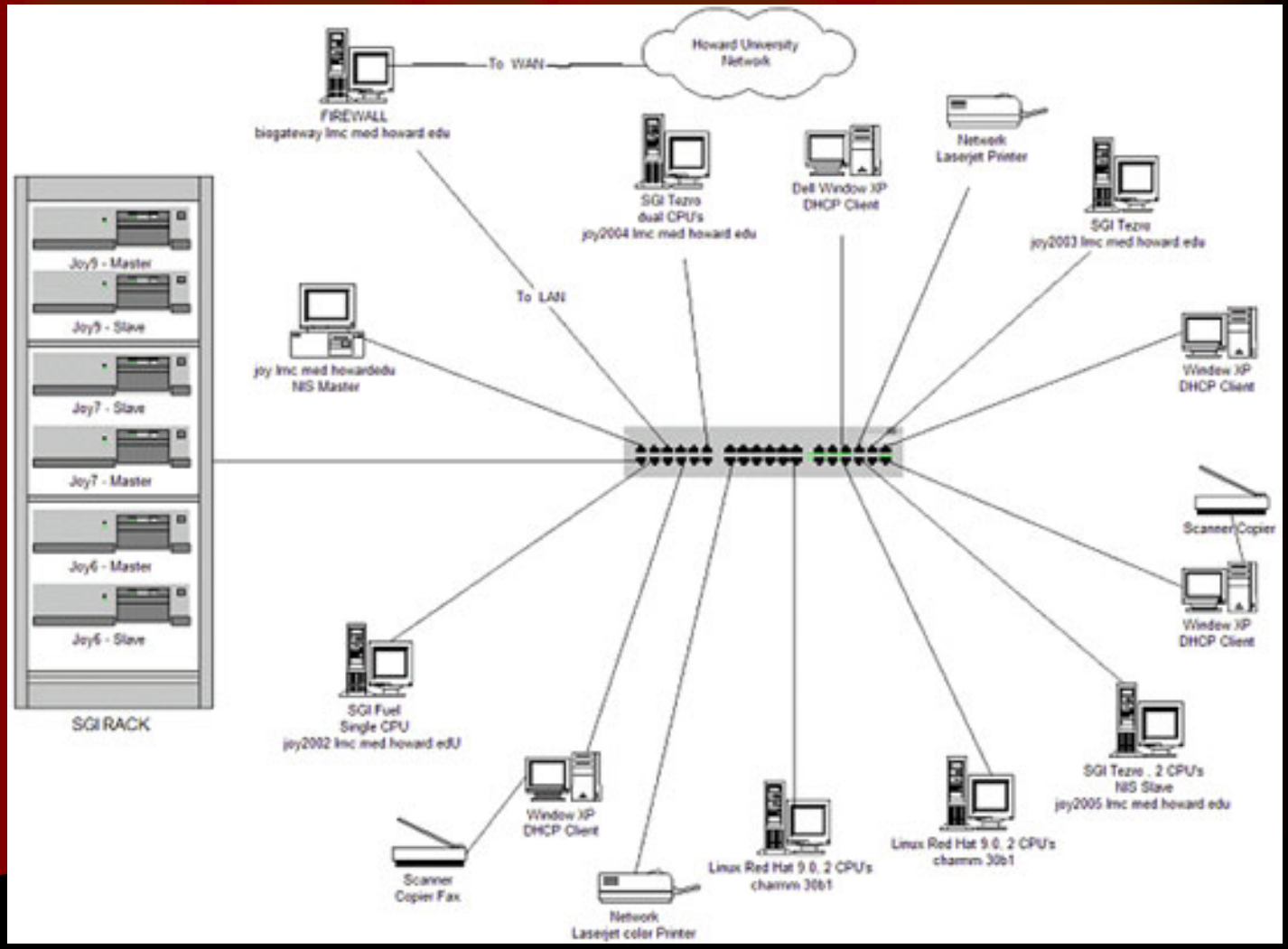
Administración de la Demanda



Actividad del Sistema



Cluster en Red



Programación y Librerías

- Compiladores
- GNU C , GNU F77/F90
- C/C++
- NAG
- Otras herramientas
- Librerias Paralelo
- Librerias Matematicas

Desarrollo Paralelo

- PADE - Parallel Aplications Development Enviroment (NITS)
- XPVM - Front-End PVM
- LAM - Local Area Multicomputer
- MOSIX

Librerias Paralelo

- ACL - Advanced Computing Lab Tools
- SILON – Scripting Interface Languages for Object- Oriented Numerics
- PAWS – Parallel Application WorkSpace
- POOMA – Parallel Object – Oriented Methods and Applications
- PETSc – Parallel scientific computing
- PLAPACK – Parallel Linear Algebra Package

Software Beowolf

ScaLAPACK

FORTRAN

Absoft FORTRAN

NAG (Numerical Algorith Group)

The Portland Group F90, C++

Parallel Tools Consortium

MAPLAB

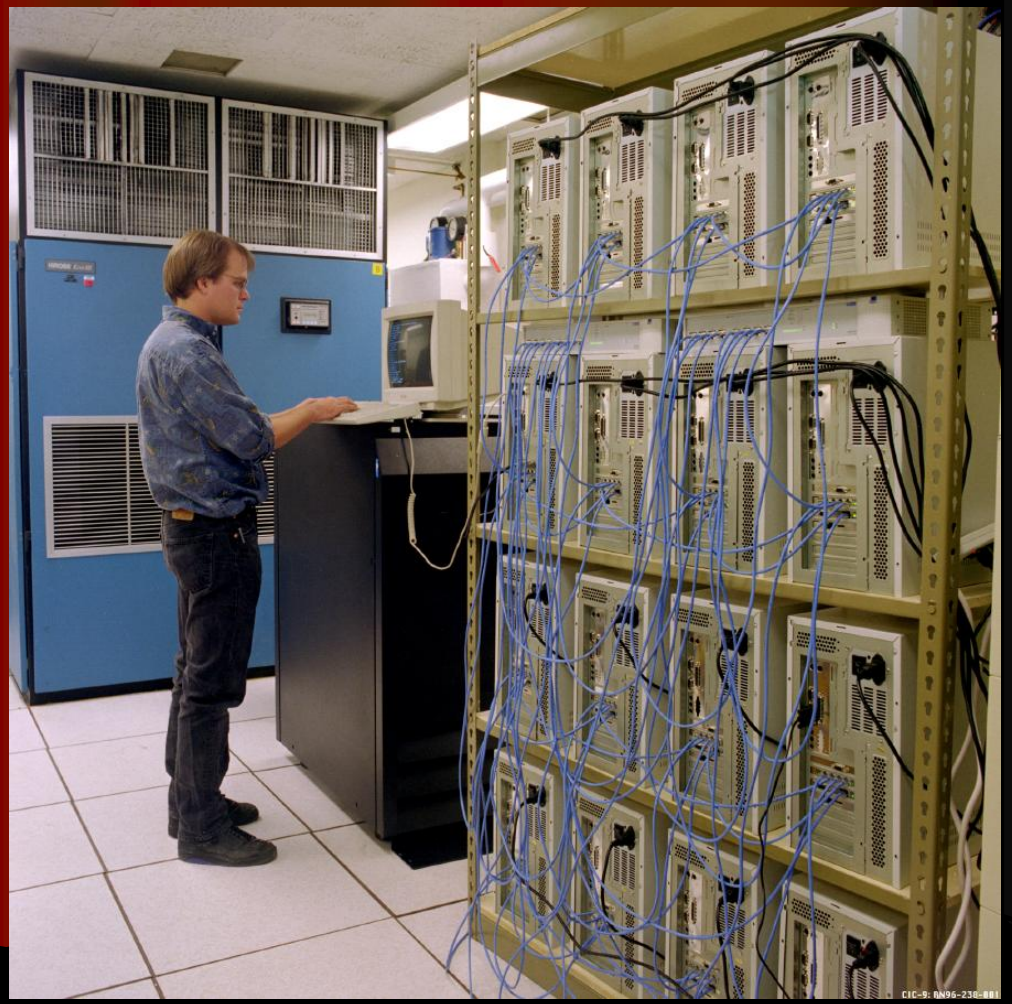
IDL

Ambiente Paralelo

- El cuarto
- Aire acondicionado
- Potencia Consumida
- UPS
- Seguridad
- Montaje

SUPER CLUSTER 512 PC

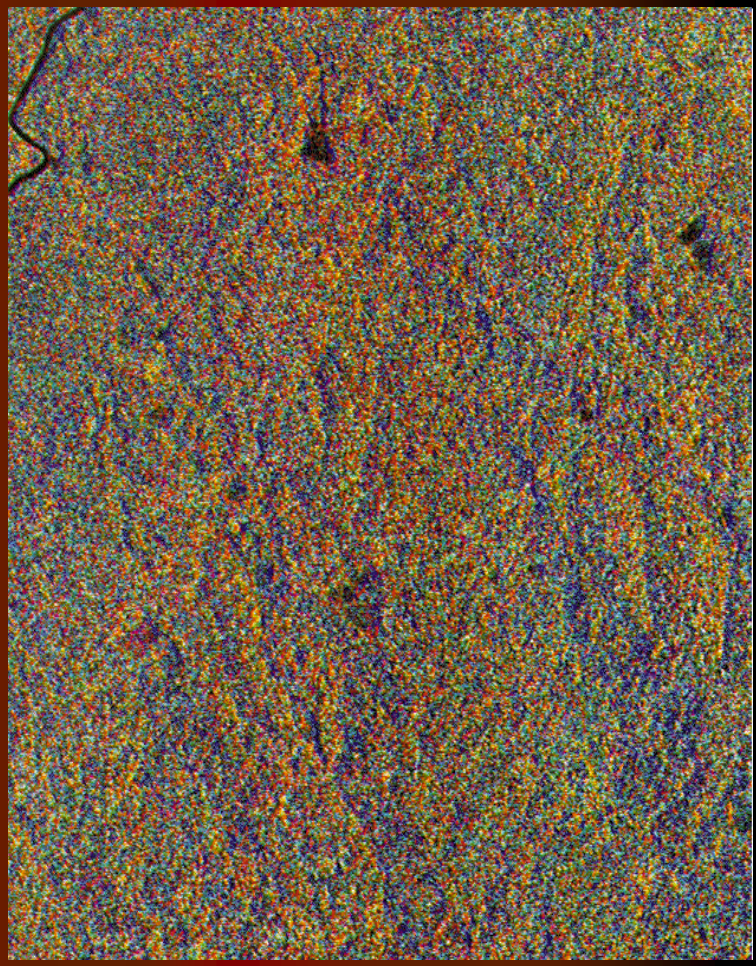
Ambiente Temperado
Falso piso
UPS
Racks



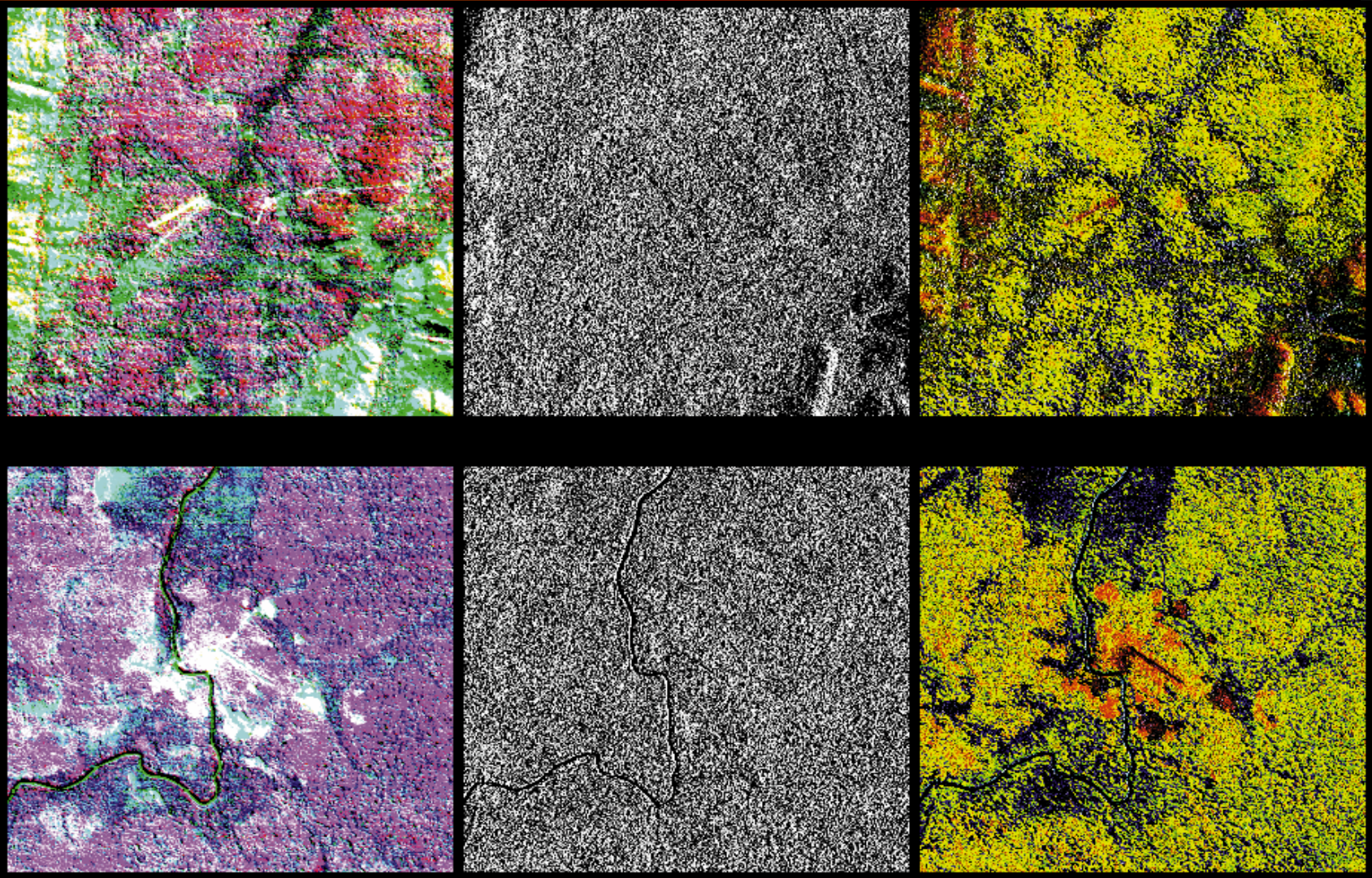
CLUSTER COMPACTO

Procesamiento

Datos SAR del satélite ERS-2
De la Agencia Espacial Europea
Procesados por equipo SUN Ultra
con dos procesadores RISC de 800 MHz.
Imagen multitemporal de zona de selva
en el Perú
250 MB cada imagen

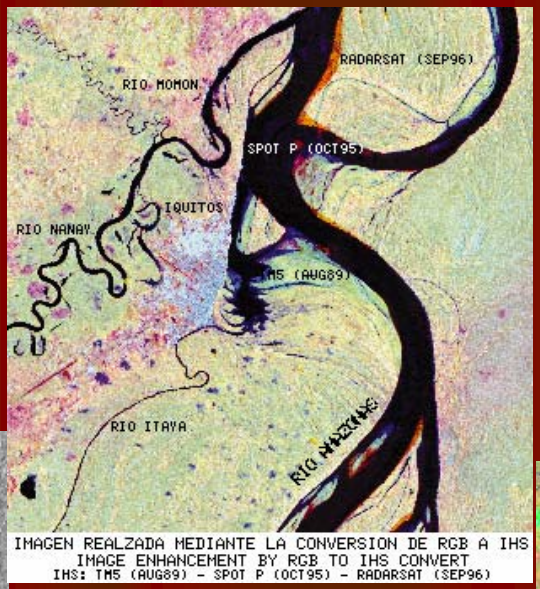


Composición Multibanda



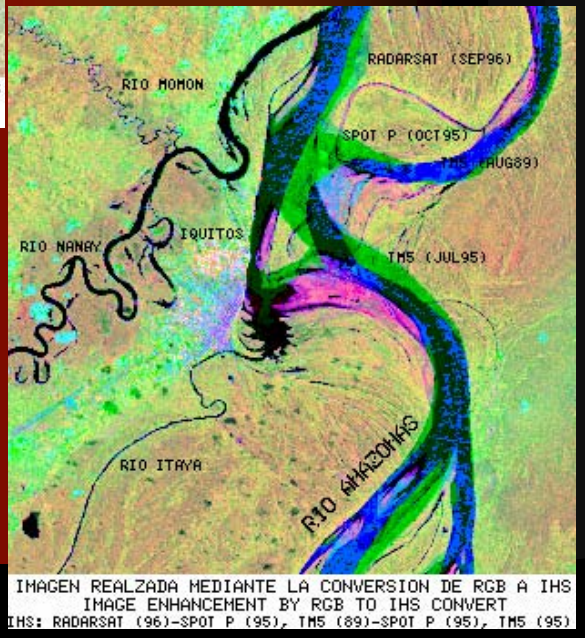
Cmbinacion de Sensores

Datos SAR
Radarsat



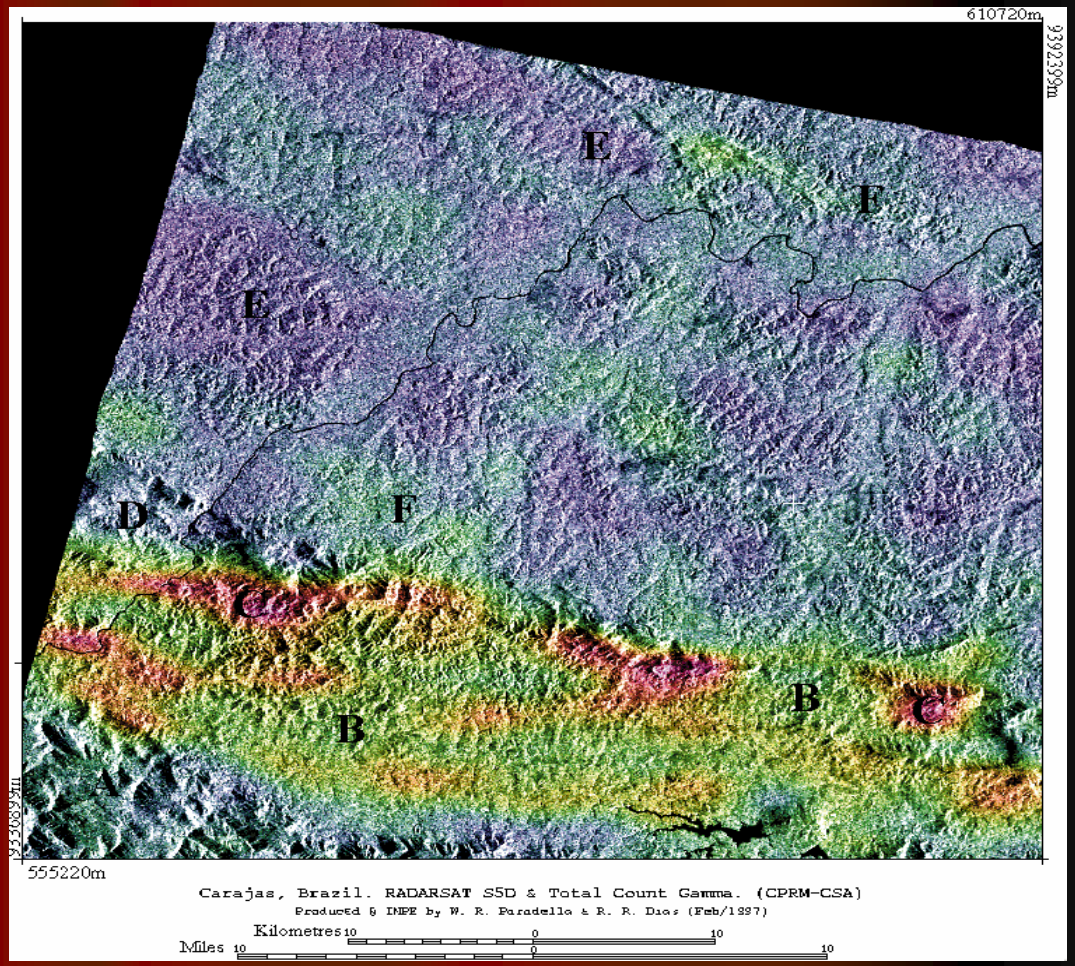
Integración de datos
SAR, SPOT-P, TMS
Formato IHS

Integración de datos
SAR, SPOT-P, TMS
Formato RGB



Proceso Multitemporal

Imagen SAR-ESA
Procesada en el INPE
Brasil por computador
Digital y Array Processor



Procesamiento · 3D

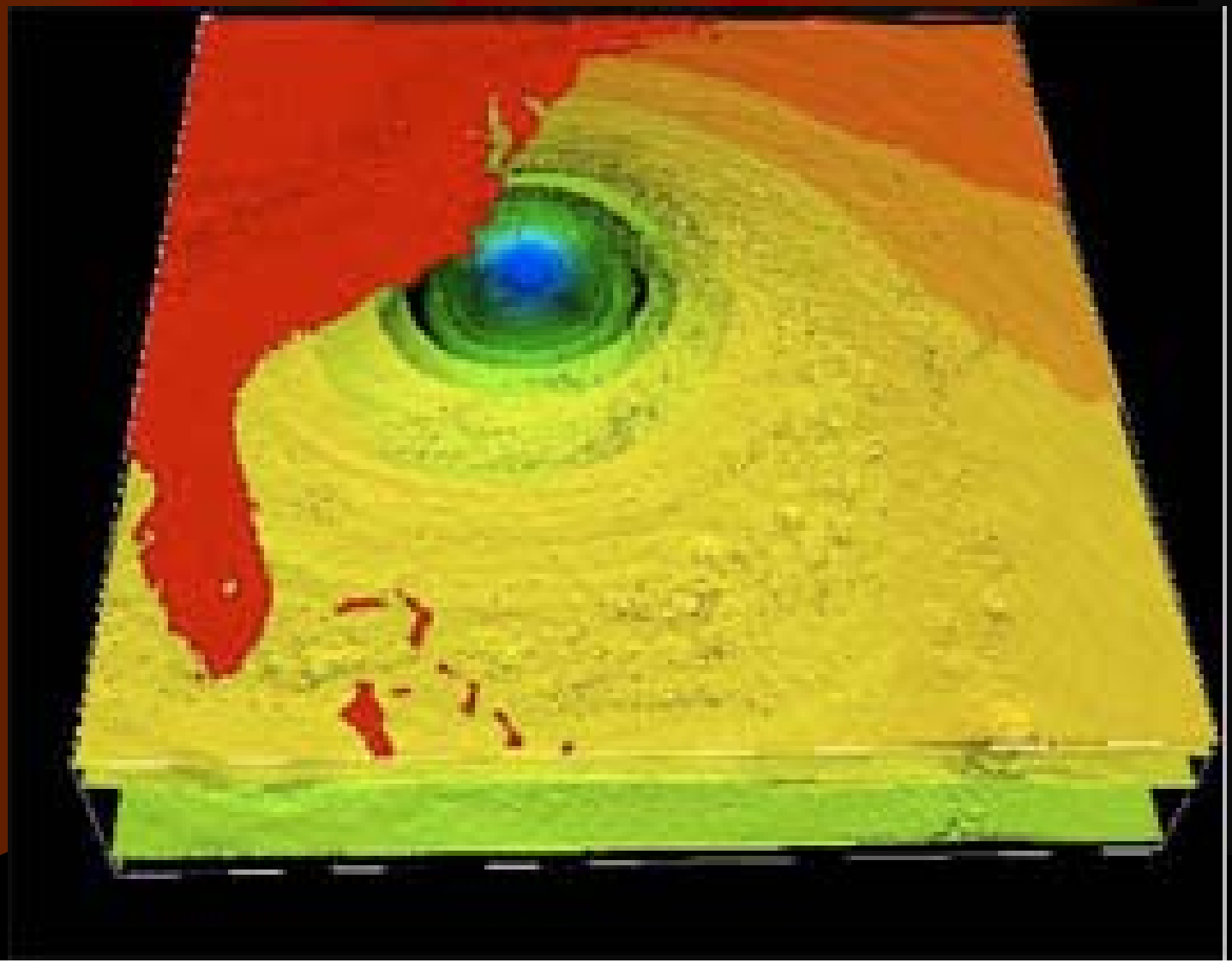
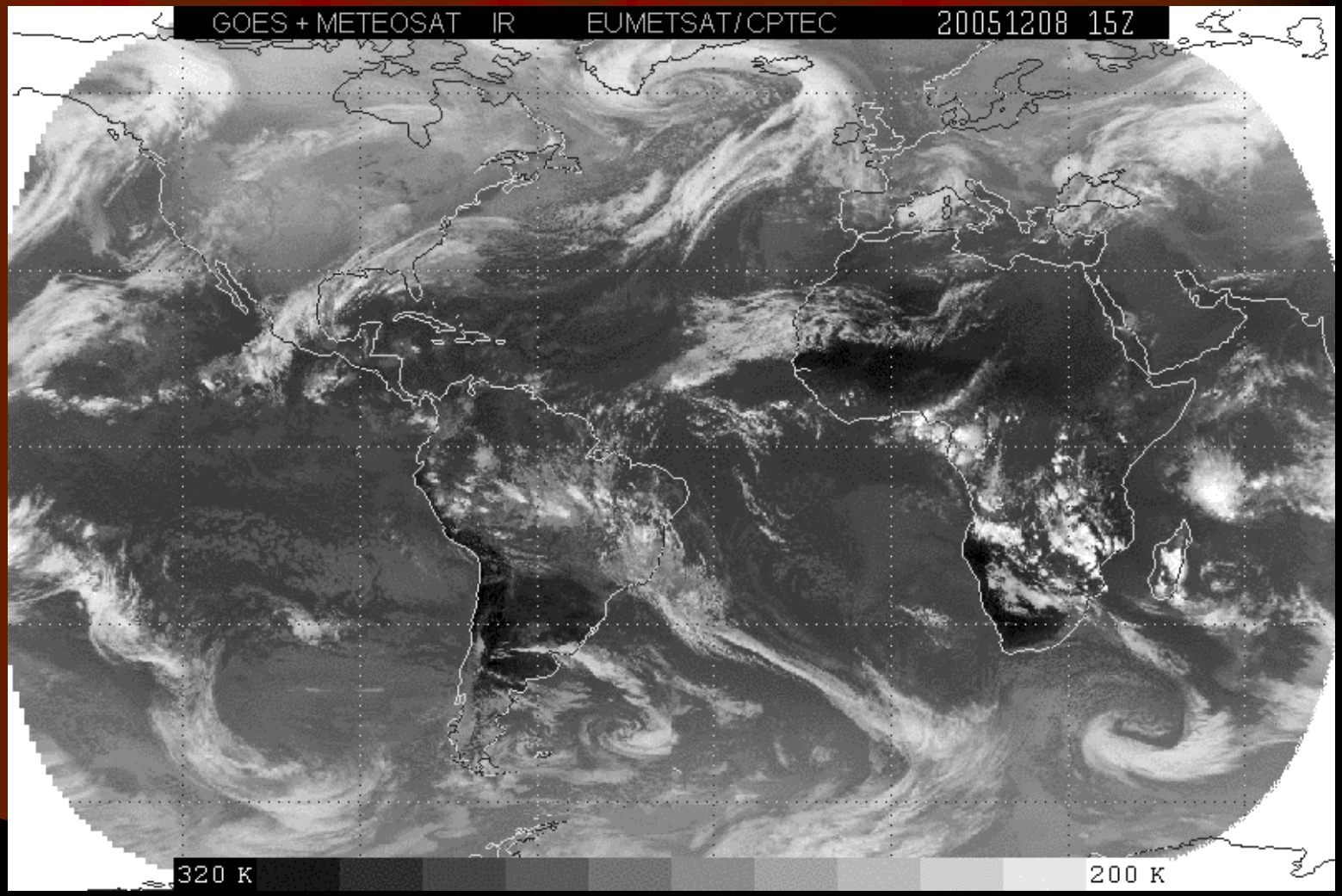


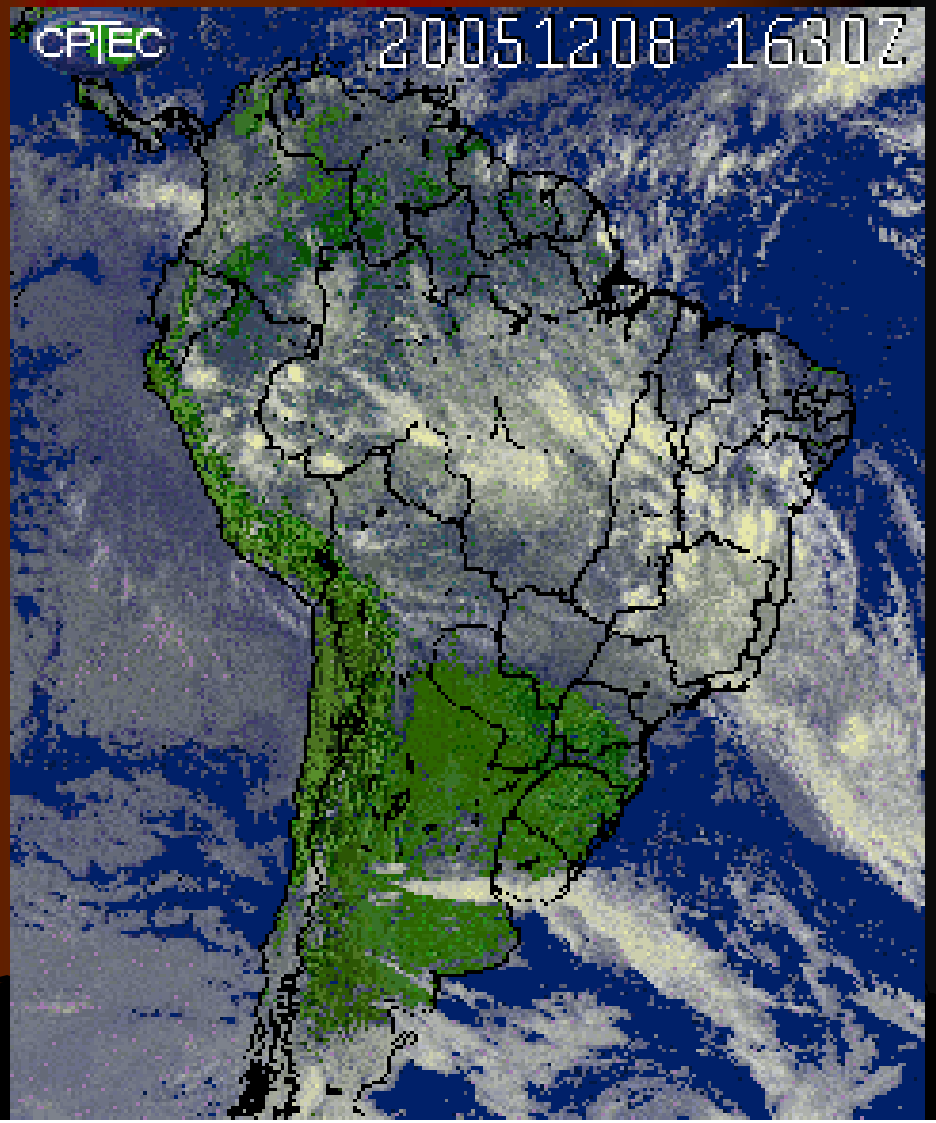
Imagen GOES



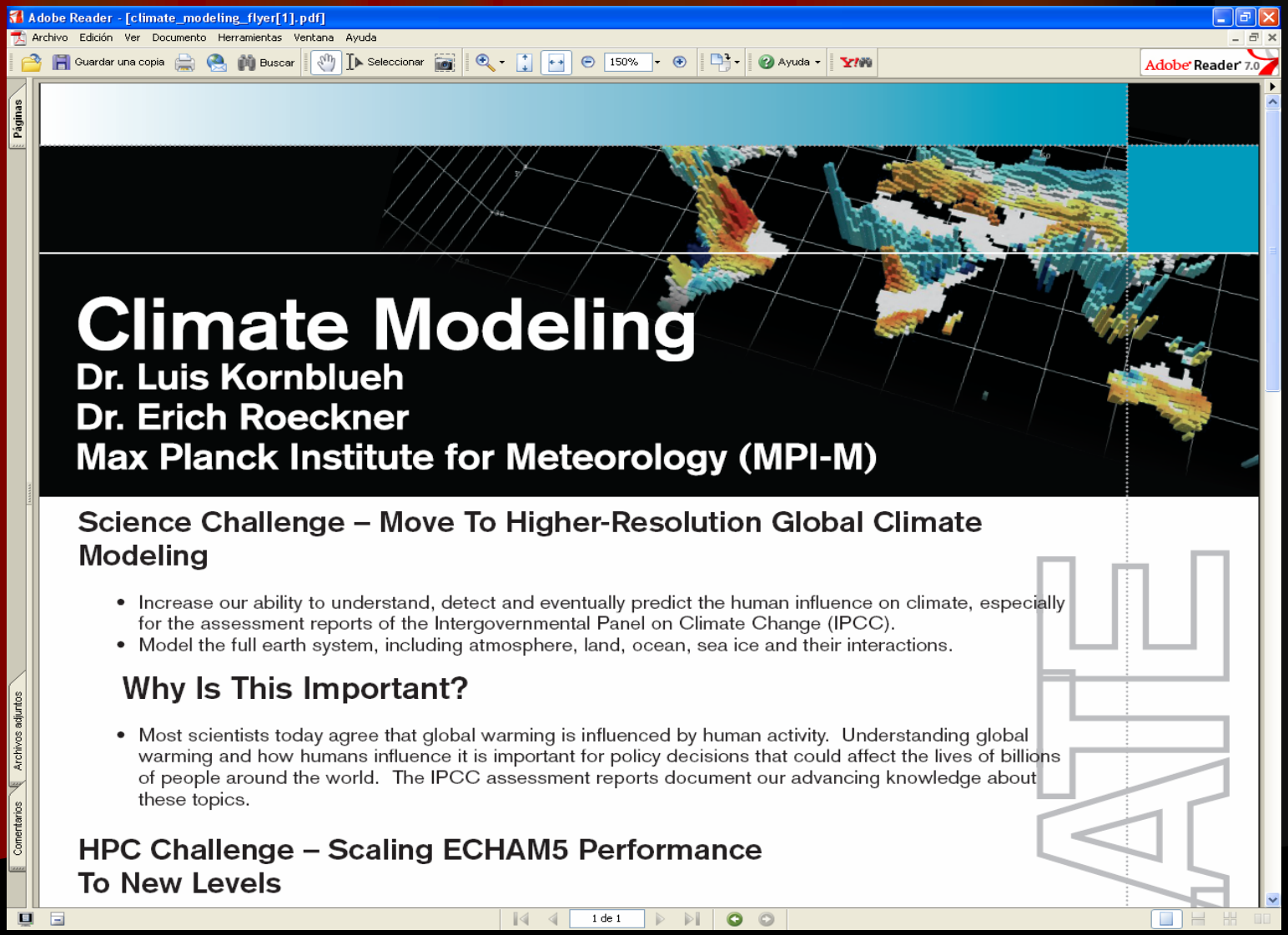
Goes Regional

Imagen utilizada por el

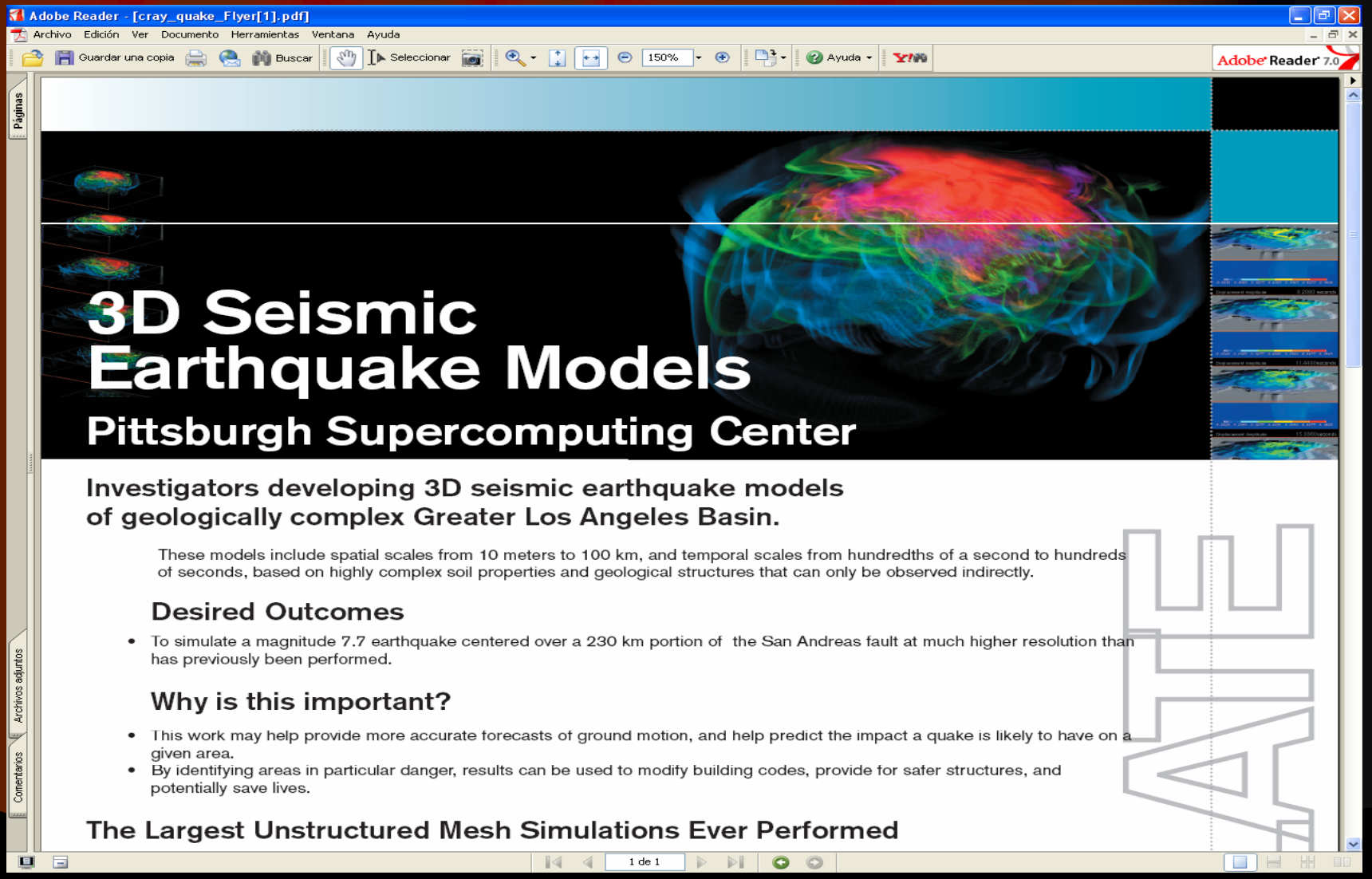
CPTEC para pronóstico



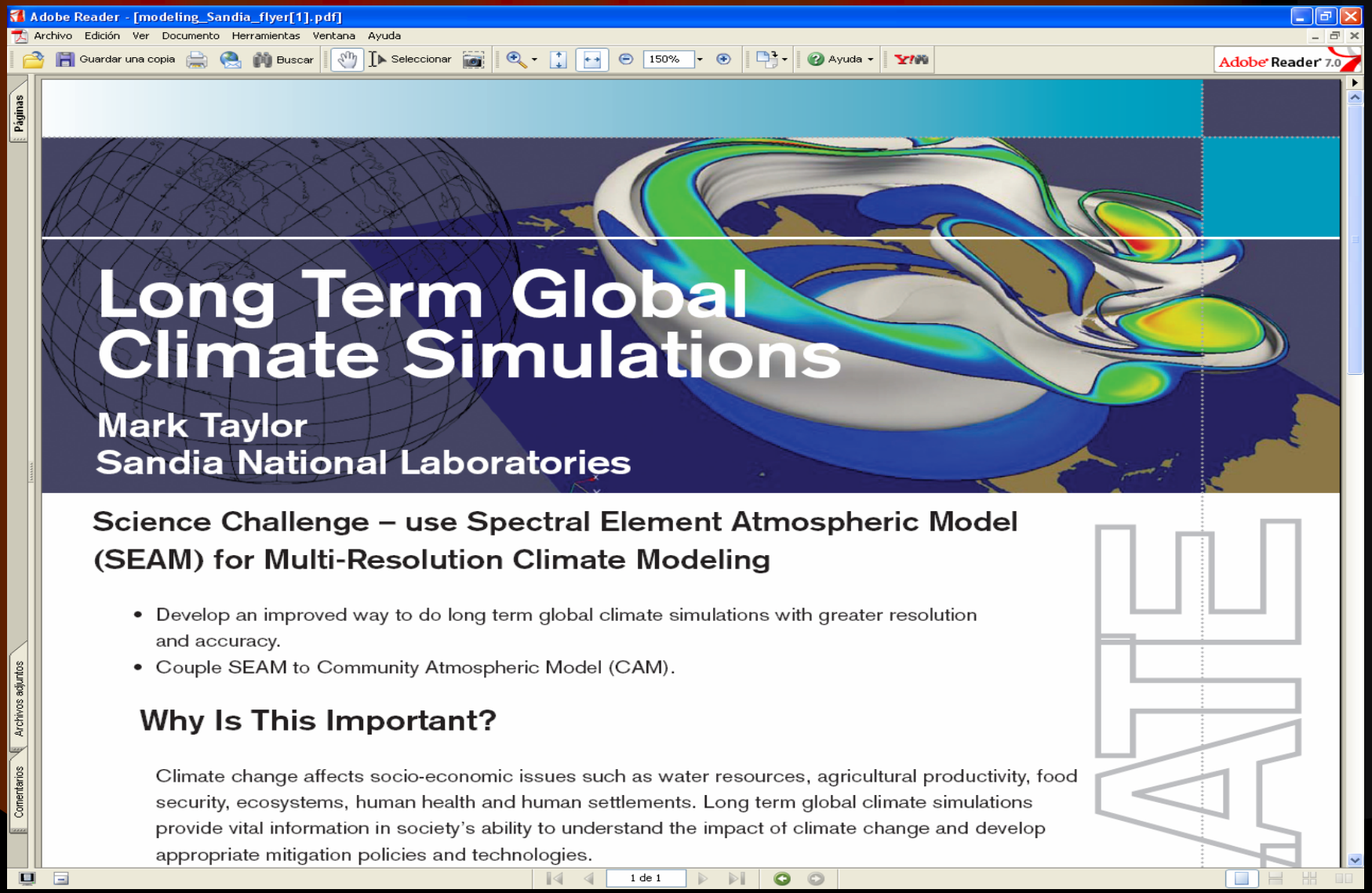
Modelado ECHAM5



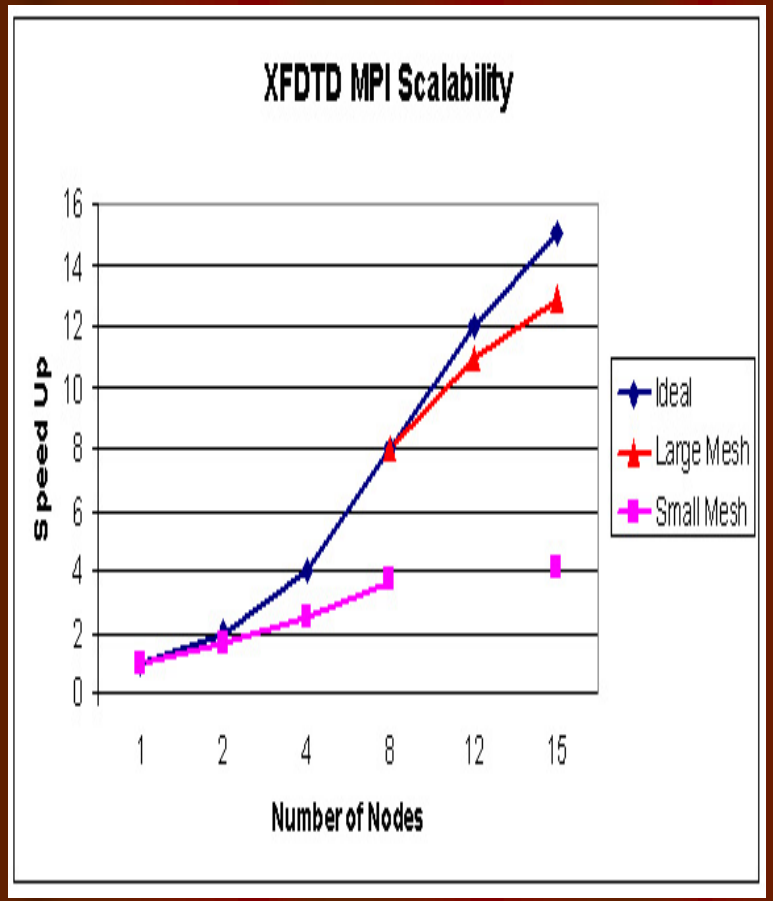
Modelado Sismico

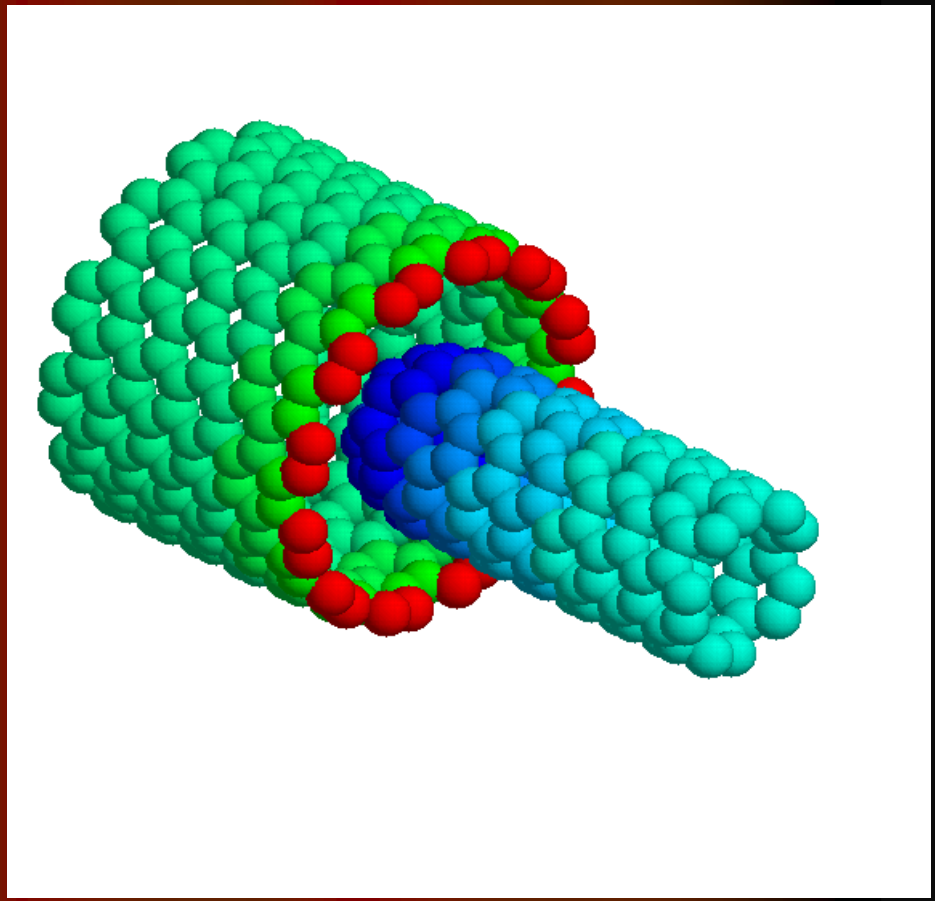


Modelado Climático



Modelos 3D


XFDTD MPI Scalability
Speed Up
16
14
12
10
8
6
4
2
0
1
2
4
8
12
15
Number of Nodes
Ideal
Large Mesh
Small Mesh




BUNYIP